УДК 378.146:51:681.3

КОМПЬЮТЕРНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ КАЧЕСТВА ИНЖЕНЕРНОГО ОБРАЗОВАНИЯ. ТЕКУЩИЙ КОНТРОЛЬ МАТЕМАТИЧЕСКИХ ЗНАНИЙ

В.П. Арефьев, А.А. Михальчук

Томский политехнический университет
E-mail: vpa@ido.tpu.edu.ru

Проведен компьютерный статистический анализ результатов текущего контроля математических знаний. Оценен вклад элитной составляющей технического образования. Дан сравнительный анализ результатов входного контроля математических знаний с результатами текущего контроля в целом и в зависимости от разных форм обучения в вузе в частности. Сделан вывод о статистически значимых различиях этих результатов. Обсуждены причины выявленных существенных различий. Рассмотрена корреляционная зависимость результатов текущего контроля от результатов входного контроля математических знаний.

Введение

Главной проблемой современного инженерного образования является проблема его качества [1, 2]. Включение России в Болонский процесс актуализирует задачу обеспечения качества высшего образования и выделяет одной из важных задач ближайшего времени разработку и использования методик оценки качества образования, сопоставимых с общеевропейскими. В частности, необходимо создание эффективной системы контроля качества знаний [3, 4]. Таким образом, если в рамках зачетно (кредитно) – рейтинговой системы оценки и контроля качества освоения образовательных программ зачетная единица оценивает объем образовательных программ, то рейтинг позволяет количественно оценить качество освоения образовательных программ [5–7].

В Томском политехническом университете (ТПУ) наряду с входным контролем (**ВК**) математических знаний школьной программы (на основе аудиторной контрольной работы с проверкой ее преподавателями) [8] применяется также в течение семестра рейтинговая система оценки математических знаний студента: проводится тематический контроль (проверка домашних индивидуальных заданий и соответствующих аудиторных контрольных работ), модульный контроль теории (коллоквиумы и проверка рефератов по самостоятельно разобранным темам), суммарный промежуточный аттестационный (**АТТ**) и суммарный осенний семестровый (по 1000-бальной рейтинговой системе) экзаменационный (**ЭКЗ**) контроль математических знаний. В связи с этим представляет интерес сравнение результатов входного контроля с соответствующими результатами текущего контроля.

Статистический метод

Сравнение результатов входного и текущего контроля проведено посредством методов статистического анализа. Все числовые результаты приведены к единой 5-бальной шкале (делением результата на соответствующий максимальный результат и умножением на пять). Созданная таким образом в MS Excel база данных использовалась далее в пакете Statistica 6.0 для статистического анализа данных [9]: вычисления числовых характеристик (объем выборки N, среднее m, стандартное отклонение σ, асимметрия A и эксцесс E, а также стандартные ошибки (Std. Err.) A и E), применения параметрического t-критерия Стьюдента для сравнения средних двух независимых или зависимых выборок (для проверки нормальности распределения использованы оценки асимметрии A и эксцесса E, критерии Пирсона и Колмогорова-Смирнова), а также непараметрических критериев: Манна-Уитни, Колмогорова-Смирнова для независимых выборок или критерий знаков и критерий Вилкоксона для зависимых выборок. В случае сравнения средних более чем двух выборок применялся однофакторный дисперсионный анализ. Для оценки корреляционной зависимости результатов текущего контроля от результатов входного контроля математических знаний вычисляется коэффициент корреляции Пирсона и используется диаграмма рассеяния.

Результаты и их анализ

В ЭКЗ набора 2004 г. участвовали 1930 студентов первого курса ТПУ: институт геологии и нефтегазового дела (ИГНД), электротехнический институт (ЭЛТИ); физико-технический (ФТФ), электрофизический (ЭФФ), машиностроительный (МСФ), химико-технологический (ХТФ), теплоэнергетический (ТЭФ), автоматики и вычислительной техники (АВТФ), естественных наук и математики (ЕНМФ) факультеты. Выборку **ЭКЗ** можно разбить на выборку **ЭКЗ**$_{э}$ объема 113 (≈6 % от общего объема) результатов контроля математических знаний студентов, отделенных в 2004 г. в группы элитного технического образования (ЭТО), развитие которого в последнее время приобретает актуальность [1, 5, 10], и всех остальных **ЭКЗ**$_{нэ}$ объема 1817.

Числовые характеристики выборок приведены в табл. 1. Средний балл **ЭКЗ**$_{э}$ (выборка ЭТО) составил 3,735 по 5-бальной шкале, что почти на балл выше среднего балла 2,828 выборки **ЭКЗ**$_{нэ}$.

Таблица 1. *Числовые характеристики выборок* ***ЭКЗ, ЭКЗ***$_{нэ}$ *и* ***ЭКЗ***$_{э}$

Выборка	N	m	σ	A	Std. Err. A	E	Std. Err. E
ЭКЗ	1930	2,882	1,231	–0,510	0,056	–0,622	0,111
ЭКЗ$_{нэ}$	1817	2,828	1,242	–0,442	0,057	–0,707	0,115
ЭКЗ$_{э}$	113	3,735	0,529	0,178	0,227	–0,504	0,451

Выборка **ЭКЗ**$_э$ распределена по нормальному закону – ***A*** и ***E*** по модулю не превышают утроенных соответствующих стандартных ошибок в отличие от выборок **ЭКЗ** и **ЭКЗ**$_{нэ}$, имеющих значимые отклонения от нормального закона как по асимметрии ***A*** вправо, так и по туповершинности ***E***. Проверка нормальности распределения выборок с помощью χ^2-критерия Пирсона и критерия Колмогорова-Смирнова дает высоко значимые отличия (уровень значимости α<0,001). Применение непараметрических критериев сравнения средних двух независимых выборок приводит к выводу о не значимом (α>0,10) различии в средних баллах 2,882 для **ЭКЗ** и 2,828 для **ЭКЗ**$_{нэ}$. Таким образом, вклад выборки **ЭКЗ**$_э$ в **ЭКЗ** является не значимым.

Для оценки значимости результатов **ЭКЗ**$_э$ рассмотрим совместное распределение выборок **ЭКЗ**, **ЭКЗ**$_{нэ}$ и **ЭКЗ**$_э$, отраженное на рис. 1, где для простоты восприятия использована неравномерная 5-бальная шкала: [0; 2,5] – «неуд»; (2,5; 3,5] – «удовл»; (3,5; 4,5] – «хор» и (4,5; 5] – «отл». Совместное распределение выборок **ЭКЗ**$_{нэ}$ и **ЭКЗ** визуально подтверждает их однородность. Как следует из рис. 1, только 8 студентов ЭТО получили в итоге семестрового обучения «отл», а больше половины – «удовл»; в тоже время число получивших «отл» в **ЭКЗ**$_{нэ}$ на порядок выше (101 студент). Это свидетельствует об остроте проблемы отбора студентов в ЭТО.

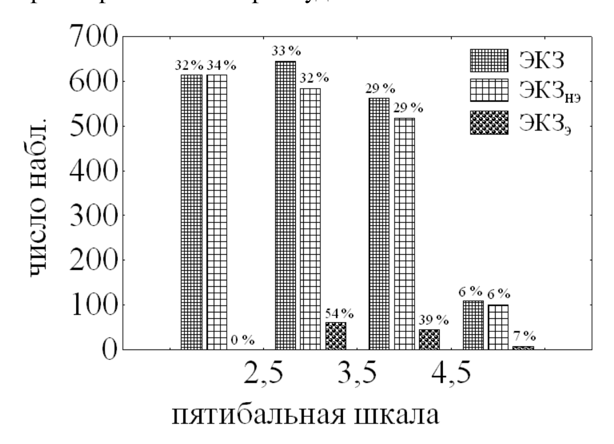


Рис. 1. *Составная гистограмма результатов **ЭКЗ, ЭКЗ**$_{нэ}$ и **ЭКЗ**$_э$*

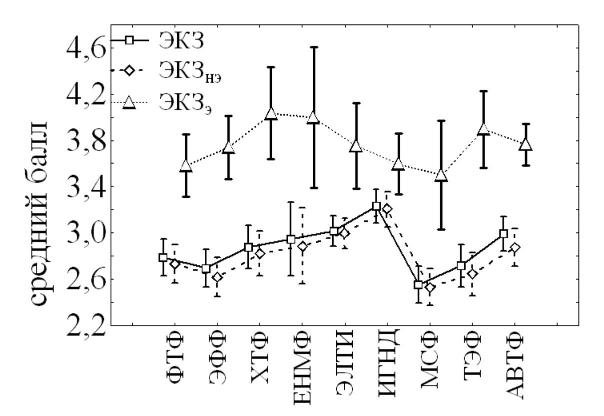


Рис. 2. *Линейные графики среднего балла с 95 % доверительным интервалом результатов **ЭКЗ, ЭКЗ**$_{нэ}$ и **ЭКЗ**$_э$ по факультетам*

Значения средних баллов результатов **ЭКЗ**, **ЭКЗ**$_{нэ}$ и **ЭКЗ**$_э$ по факультетам с указанием 95 % доверительного интервала приведены на рис. 2. Совместное распределение выборок **ЭКЗ**$_{нэ}$ и **ЭКЗ** по факультетам визуально еще раз подтверждает их однородность.

Однородность независимых выборок (факультетов) была оценена посредством однофакторного дисперсионного анализа применительно к выборке **ЭКЗ**. Применение ***F***-критерия, обобщающего ***t***-критерий на число выборок больше двух, привело к выводу о неоднородности сравниваемых факультетов. Значимость различий между средними баллами ***m*** по факультетам оценивалась с помощью апостериорного критерия наименьших значений разности, эквивалентного ***t***-критерию для числа независимых выборок больше двух. Результаты сравнения средних баллов m по факультетам отражены в табл. 2.

Таблица 2. *Уровни значимости α различий средних баллов m в независимых выборках по факультетам*

ФАК	**ФТФ**	**ЭФФ**	**ХТФ**	**ЕНМФ**	**ЭЛТИ**	**ИГНД**	**МСФ**	**ТЭФ**	**АВТФ**
ФТФ		0,418	0,480	0,380	0,029	0,000	0,038	0,550	0,066
ЭФФ	0,418		0,151	0,165	0,002	0,000	0,210	0,873	0,008
ХТФ	0,480	0,151		0,712	0,237	0,004	0,010	0,225	0,351
ЕНМФ	0,380	0,165	0,712		0,691	0,112	0,029	0,213	0,800
ЭЛТИ	0,029	0,002	0,237	0,691		0,033	0,000	0,008	0,811
ИГНД	0,000	0,000	0,004	0,112	0,033		0,000	0,000	0,026
МСФ	0,038	0,210	0,010	0,029	0,000	0,000		0,181	0,000
ТЭФ	0,550	0,873	0,225	0,213	0,008	0,000	0,181		0,021
АВТФ	0,066	0,008	0,351	0,800	0,811	0,026	0,000	0,021	
m	2,788	2,695	2,876	2,946	3,015	3,228	2,551	2,715	2,991

Тональность заливок ячеек табл. 2 соответствует степени значимости различий средних ***m*** в независимых выборках по факультетам в выборке **ЭКЗ**: не значимые (α≥0,100), слабо значимые (0,100>α≥0,050), значимые (0,050>α≥0,010), сильно значимые (0,010>α≥0,001), высоко значимые (0,001>α). Как следует из табл. 2, имеются всевозможные степени значимости различий между факультетами.

Для проведения сравнительного анализа входного и текущего контроля качества математических знаний сформируем выборку студентов трех факультетов (ФТФ + ЭФФ + ХТФ), участвовавших как во входном контроле (**вк**), так и в текущем: аттестационном октябрьском (**атт**$_1$), аттестационном ноябрьском (**атт**$_2$) и осеннем семестровом экзаменационном (**экз**) контроле математических знаний. Числовые характеристики выборок приведены в табл. 3.

Таблица 3. *Числовые характеристики выборок входного и текущего контроля*

Выборка	***N***	***m***	σ	***A***	Std. Err. ***A***	***E***	Std. Err. ***E***
ВК	1175	2,216	1,406	0,127	0,071	**−0,972**	**0,143**
вк	561	2,157	1,371	0,191	0,103	**−0,951**	**0,206**
атт$_1$	561	2,763	1,227	−0,278	0,103	**−0,724**	**0,206**
атт$_2$	561	2,755	1,151	−0,223	0,103	**−0,695**	**0,206**
экз	561	2,828	1,171	**−0,517**	**0,103**	−0,421	0,206
ЭКЗ	1930	2,882	1,231	**−0,510**	**0,056**	**−0,622**	**0,111**

Сначала было проведено сравнение результатов входного контроля **ВК** и **вк**, а также экзаменационного контроля **ЭКЗ** и **экз**. Результаты входного и экзаменационного контроля отображены на рис. 3 и 4. Для визуальной оценки сходства наблюдаемых распределений (гистограммы) с теоретическим распределением по нормальному закону (соответствующие кривые) использована равномерная 5-бальная шкала.

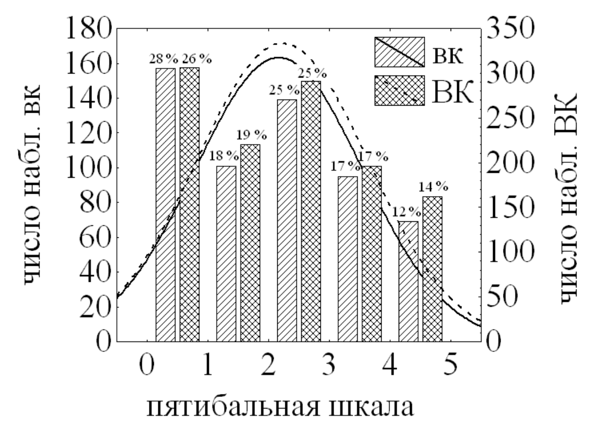


*Рис. 3. Составная гистограмма результатов **вк** и **ВК***

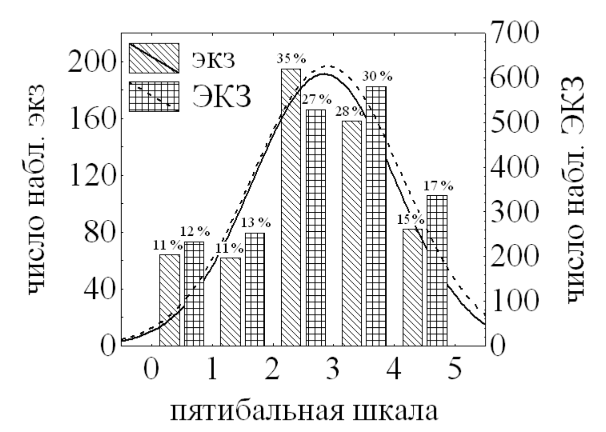


*Рис. 4. Составная гистограмма результатов **экз** и **ЭКЗ***

Уровень значимости различия средних результатов **ВК** и **вк**, а также **ЭКЗ** и **экз** определялся с помощью статистических критериев. Сравнение проведено по аналогии со сравнением выборок $\textbf{ЭКЗ}_{\text{НЭ}}$ и **ЭКЗ**. Статистические критерии сравнения средних двух независимых выборок дали не значимые (α>0,10) различия как в средних баллах результатов входного контроля 2,216 для **ВК** и 2,157 для **вк**, так и в средних баллах результатов экзаменационного контроля 2,882 для **ЭКЗ** и 2,828 для **экз**. Таким образом, далее можно ограничиться рассмотрением выборок **вк** и **экз**.

Значимость различий между средними ***m*** выборок **вк**, $\textbf{атт}_1$, $\textbf{атт}_2$ и **экз** оценивалась с помощью дисперсионного анализа с повторными измерениями, обобщающего ***t***-критерий для двух зависимых выборок на число выборок больше двух. Результаты сравнения средних выборок входного (**вк**) и текущего ($\textbf{атт}_1$, $\textbf{атт}_2$, **экз**) контроля отражены в табл. 4.

Как следует из табл. 4, выборки текущего контроля $\textbf{атт}_1$ и $\textbf{атт}_2$ различаются не значимо, в то время как выборка **вк** входного контроля высоко значимо отличается от всех выборок текущего контроля.

Таблица 4. *Уровни значимости α различий средних **m** в зависимых выборках входного и текущего контроля*

выборка	**вк**	$\textbf{атт}_1$	$\textbf{атт}_2$	**экз**
вк		0,000	0,000	0,000
$\textbf{атт}_1$	0,000		0,792	0,110
$\textbf{атт}_2$	0,000	0,792		0,024
экз	0,000	0,110	0,024	

Совместное распределение выборок **вк**, $\textbf{атт}_1$, $\textbf{атт}_2$ и **экз** отображенно на рис. 5, где для простоты восприятия использована неравномерная 5-бальная шкала. Визуально очевидна динамика улучшения качества уровня математических знаний: монотонное убывание «неуд» и монотонный рост «удовл» и «хор». Хуже обстоят дела с «отл», что можно объяснить постоянным уменьшением в последнее десятилетие объема аудиторных часов изучения высшей математики, чрезмерное акцентирование на самостоятельной работе и, как следствие, повышение интенсивности изучения математики настолько, что у преподавателя уже не хватает времени для должной работы с лучшими студентами.

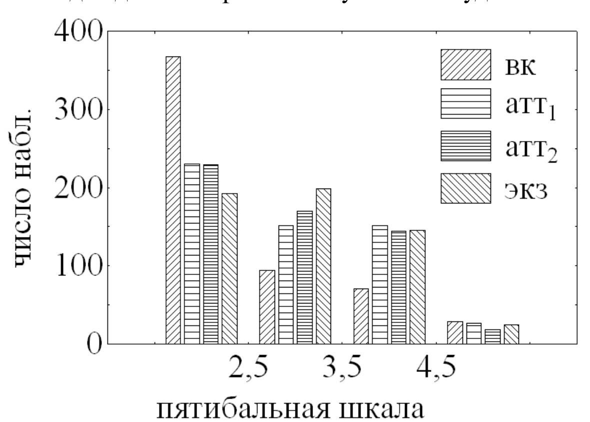


*Рис. 5. Составная гистограмма результатов **вк**, $\textbf{атт}_1$, $\textbf{атт}_2$ и **экз***

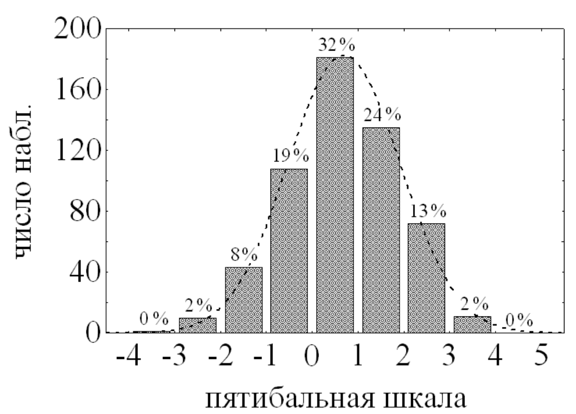


*Рис. 6. Гистограмма распределения разности **экз–вк***

Рассмотрим подробнее наиболее контрастные выборки результатов входного **вк** и суммарного семестрового **экз** контроля математических знаний. В связи с проверкой применения ***t***-критерия Стью-

дента для зависимых выборок необходимо проверить выборку поэлементной разности **экз-вк** (рис. 6) на нормальность распределения. Проверка нормальности распределения выборки разности **экз-вк** с помощью коэффициентов асимметрии ***A*** и эксцесса ***E***, χ^2-критерия Пирсона и критерия Колмогорова-Смирнова дает не значимые отличия распределения разности **экз-вк** от нормального закона. Можно отметить, что результаты разности **экз-вк** представляют и самостоятельный интерес. Как следует из рис. 6, почти 3/4 студентов (71 %) повысили качество математических знаний в результате обучения в первом семестре. Это отражает динамику улучшения качества математических знаний: рост среднего балла с 2,157 (для **вк**) до 2,828 (для **экз**). Но, с другой стороны, есть еще и 29 % студентов, которые понизили качество своих математических знаний. При этом, например, 2 % студентов имеют результаты **экз** ниже соответствующих результатов **вк** более чем на два балла по 5-бальной шкале. Вероятно, это случайно поступившие в ТПУ. Не последней причиной понижения качества математических знаний в остальных 27 % является возросшая интенсивность обучения математики в ТПУ.

Представляет интерес исследование результатов разности **экз-вк** в зависимости от формы обучения (ФО): бюджетная (Б), контрактная (К) и целевой набор (Ц). Числовые характеристики выборок **экз** и **вк** для разных ФО приведены в табл. 5.

Таблица 5. *Числовые характеристики выборок **экз** и **вк** для разных ФО*

	вк	вк	экз	вк	экз
ФО	***N***	***m***	***m***	***σ***	***σ***
Ц	114	1,754	2,624	1,196	1,101
К	100	1,658	2,403	1,144	1,212
Б	347	2,433	3,017	1,414	1,141
Б+Ц+К	561	2,157	2,828	1,371	1,171

Из сравнения средних баллов следует, что наиболее низкие средние баллы как в **вк**, так и в **экз** в контрактной форме обучения.

Результаты разности **экз-вк** в зависимости от ФО представлены на круговой гистограмме (рис. 7). Как следует из рис. 7, худшая динамика улучшения качества математических знаний присуща бюджетной форме обучения.

Рассмотрим вопрос о статистической зависимости результатов **вк** и **экз**, в частности для разных ФО. Корреляция между результатами **вк** и **экз** может характеризоваться прежде всего коэффициентом парной корреляции Пирсона ***r***, отражающим степень линейной связи межу результатами **вк** и **экз**. В данном случае $r \approx 0,5$ для разных ФО, т. е. далек от единицы, а, следовательно, положительная корреляционная зависимость результатов **вк** и **экз** далека от линейной. Более того, согласно диаграмме рассеяния (рис. 8), облако значений пары (**вк**, **экз**) достаточно сильно распылено около прямой регрессии, что характеризует корреляционную зависимость результатов **вк** и **экз** как не высокую.

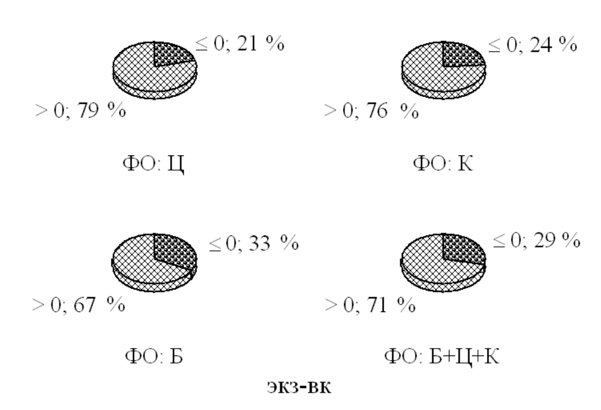


Рис. 7. *Круговая гистограмма распределения разности **экз–вк** в зависимости от форм обучения*

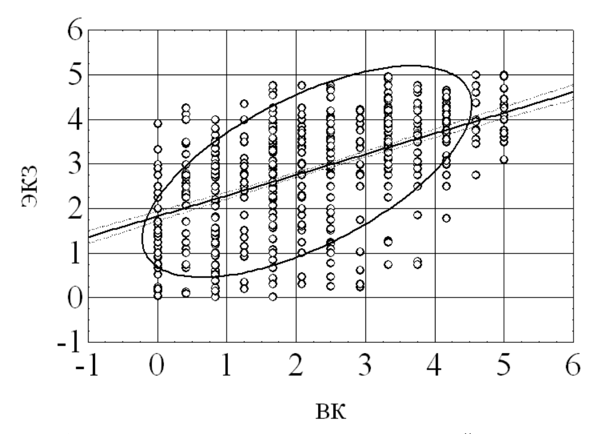


Рис. 8. *Диаграмма рассеяния корреляционной зависимости результатов **вк** и **экз***

Возвращаясь к полной выборке **ВК** и соответствующей ей выборке $\textbf{ЭКЗ}_{\textbf{ВК}}$ тех 1175 из 1930 сдававших экзамен, которые участвовали во входном контроле, можно оценить динамику улучшения качества математических знаний в зависимости от диапазонов выборки **ВК** (рис. 9). Так, например, 15 % из 26 %, набравших на **ВК** не более одного балла по 5-бальной шкале, так и не сдали экзамен на положительный результат, а среди 11 % сдавших сдали на «удовл» 8 % и на «хор» 3 %. Причем, последние 3 % – это, вероятно, формально участвовавшие, но фактически проигнорировавшие необязательный **ВК** хорошие студенты. В связи с этим предлагается придать **ВК** обязательный статус и включить его в рейтинговую систему оценки качества математических знаний. Более того, проводить обязательный **ВК** в начале каждого семестра. Таким образом, налицо хотя и скромная, но положительная динамика улучшения качества математических знаний, чего не скажешь о тех 14%, которые набрали на **ВК** более четырех баллов по 5-бальной шкале: их средний балл в **ЭКЗ** составляет 3,87. Для промежуточных диапазонов выборки **ВК** характерен разброс в обе стороны результатов в **ЭКЗ**. Так, например, 17 %, набравших на **ВК** от трех до четырех баллов по 5-бальной шкале, распределились в **ЭКЗ** при среднем балле 3,45 следующим образом: «неуд» – 3 %, «удовл» – 5 %, «хор» – 7 %,

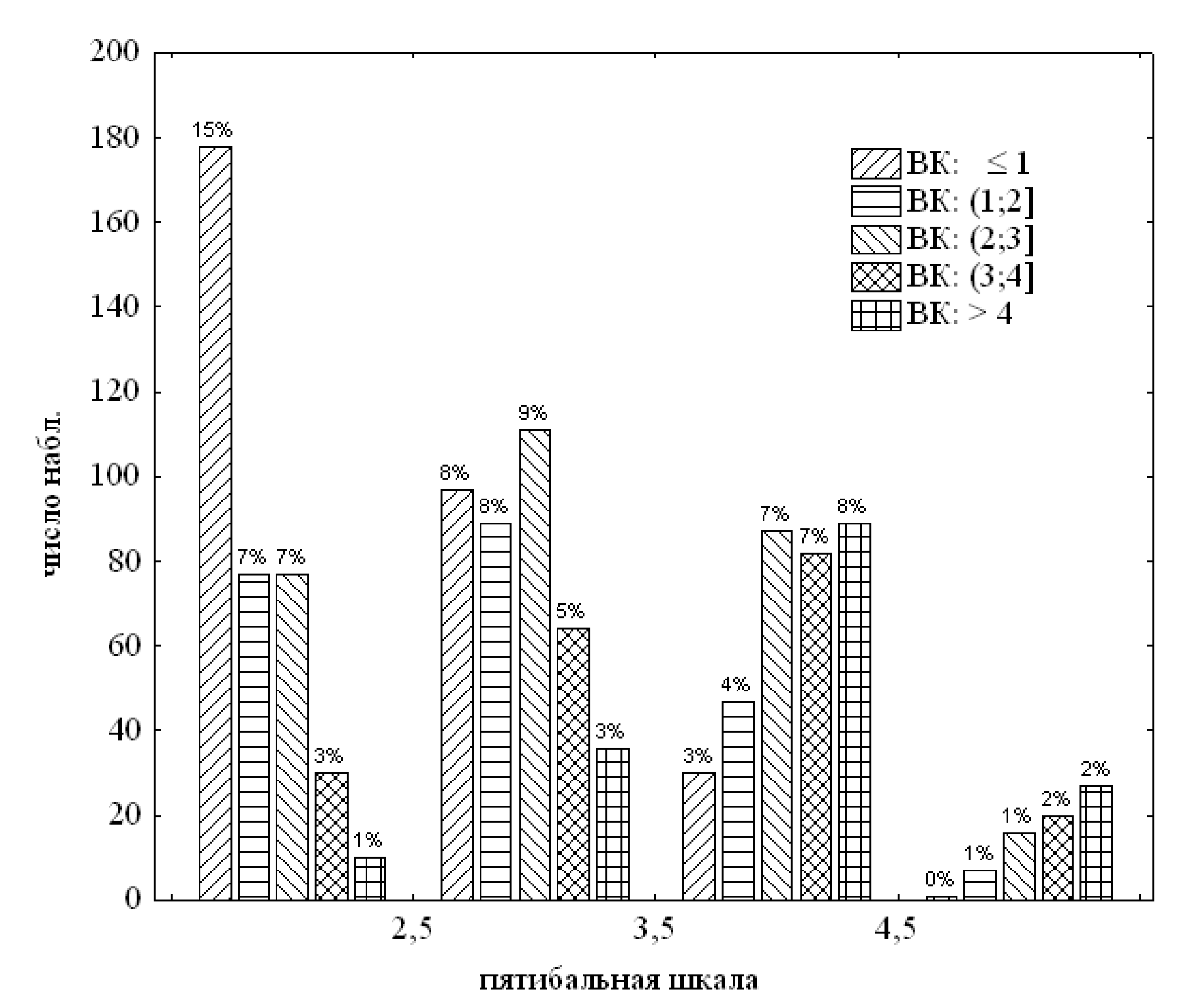


Рис. 9. *Составная гистограмма результатов* ***ЭКЗ***$_{ВК}$ *(**ВК**)*

«отл» – 2 %. Это лишний раз подчеркивает слабую корреляцию результатов входного и итогового семестрового контроля (рис. 8).

Можно заметить, что корреляция результатов входного испытания [8] и итогового семестрового контроля еще слабее (***r***≈0,42).

Выводы

1. Вклад выборки **ЭКЗ**$_Э$ (элитной составляющей) в **ЭКЗ** является статистически не значимым. Это свидетельствует об остроте проблемы отбора студентов в ЭТО.
2. Различия средних ***m*** в независимых выборках по факультетам в выборке **ЭКЗ** являются всевозможными по степени значимости.
3. Выборка **вк** результатов входного контроля высоко значимо отличается (ниже) от всех выборок результатов текущего контроля. Это отражает динамику улучшения качества математических знаний. Вместе с тем существует проблема повышения интенсивности изучения математики.
4. В контрактной форме обучения наиболее низкие средние баллы как в **вк**, так и в **экз**.
5. Для бюджетной формы обучения присуща самая низкая динамика улучшения качества математических знаний.
6. Корреляционную зависимость результатов **вк** и **экз**, равно как и **ВК** и **ЭКЗ**$_{ВК}$, можно характеризовать как не высокую.
7. Предлагается придать **ВК** обязательный статус, включить его в рейтинговую систему оценки качества математических знаний и проводить его в начале каждого семестра.
8. В целях уменьшения влияния субъективного фактора на контроль знаний и усиления корреляции **ВК** и **ЭКЗ** необходимо создание независимой, адекватной и единообразной системы контроля знаний.

СПИСОК ЛИТЕРАТУРЫ

1. Похолков Ю.П. Проблемы и основные направления совершенствования инженерного образования // Alma Mater. Вестник высшей школы. – 2003. – № 10. – С. 3–8.
2. Любимов Л. Реформа образования: благие намерения, обретения, потери // Высшее образование в России. – 2004. – № 12. – С. 8–24.
3. Сенашенко В., Ткач Г. Болонский процесс и качество образования // Alma Mater. Вестник высшей школы. – 2003. – № 8. – С. 8–14.
4. Мотова Г. Система оценки качества образования в странах СНГ и Балтии // Alma Mater. Вестник высшей школы. – 2004. – № 1. – С. 37–40.
5. Сенашенко В., Ткач Г. О ходе интеграции российской системы образования в европейское образовательное пространство // Alma Mater. Вестник высшей школы. – 2004. – № 7. – С. 13–19.
6. Давыдова Л. О показателях качества образования // Высшее образование в России. – 2004. – № 11. – С. 92–96.
7. Чучалин А., Боев О. Кредитно-рейтинговая система // Высшее образование в России. – 2004. – № 3. – С. 34–39.
8. Арефьев В.П., Михальчук А.А., Кулебакина Н.Н. Компьютерный статистический анализ качества инженерного образования. Входной контроль математических знаний. // Известия Томского политехнического университета. – 2005. – Т. 308. – № 2. – С. 201–205.
9. Боровиков В.П. STATISTICA. Искусство анализа данных на компьютере: Для профессионалов. – СПб.: Питер, 2003. – 688 с.
10. Похолков Ю.П., Вайсбурд Д.И., Чубик П.С. Элитное образование в традиционном техническом университете // Элитное техническое образование: Труды Междунар. конф. в рамках симпозиума. – Томск, 2003. – С. 6–8.